

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/10/029,656

TIME: 07:56:34

Input Set : A:\ES.txt

Output Set: N:\CRF3\01152002\J029656.raw

ENTERED

3 <110> APPLICANT: Patience, Clive
4 Oldmixon, Beth
5 Ericsson, Thomas
7 <120> TITLE OF INVENTION: Molecular Sequence of Pig Endogenous Retrovirus Receptor and
Methods of
8 Use
10 <130> FILE REFERENCE: 329579-3
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/029,656
C--> 12 <141> CURRENT FILING DATE: 2001-12-21
12 <150> PRIOR APPLICATION NUMBER: US/60/285,103
13 <151> PRIOR FILING DATE: 2001-04-20
15 <160> NUMBER OF SEQ ID NOS: 23
17 <170> SOFTWARE: PatentIn version 3.0
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 1959
21 <212> TYPE: DNA
22 <213> ORGANISM: Viral
24 <400> SEQUENCE: 1
25 atgcatccca cgttaagccg gcgccacctc ccgattcggg gtggaaagcc gaaaagactg 60
27 aaaatcccct taagcttcgc ctccatcgcg tggttcctta ctctgtcaat aactcctcaa 120
29 gttaatggta aacgccttgt gaacagtcgg aactcccata aacccttata tctcacctgg 180
31 ttacttactg actccggtac aggtattaat attaacagca ctcaagggga ggctcccttg 240
33 gggacctggt ggcctgaatt atatgtctgc ctctgatcag taatccctgg tctcaatgac 300
35 caggccacac ccccgatgt actccgtgct tacgggtttt acgtttgccc aggaccccca 360
37 aataatgaag aatattgttg aaatcctcag gatttctttt gcaagcaatg gagctgcgta 420
39 acttctaatt atgggaattg gaaatggcca gtctctcagc aagacagagt aagttactct 480
41 tttgttaaca atcctaccag ttataatcaa tttaattatg gccatgggag atggaaagat 540
43 tggcaacagc gggtaacaaa agatgtacga aataagcaaa taagctgtca ttcgttagac 600
45 ctagattact taaaataaag tttcactgaa aaaggaaaac aagaaaatat cctaaaatgg 660
47 gtacatggta tgtcttgagg aatggtatat tatggaggct cgggtaaaaca accaggctcc 720
49 attctaacta ttgcctcaa aataaaccag ctggagcctc caatggctat aggaccaaat 780
51 acggtottga cgggtcaaaag acccccaccc caaggaccag gaacatcctc taacataact 840
53 tctggatcag accccactga gtctaacagc acgactaaaa tgggggcaaa acttttttagc 900
55 ctcatccagg gagcttttca agctcttaac tccacgactc cagaggctac ctcttcttgt 960
57 tggctatgct tagcttttgg cccaccttac tatgaaggaa tggctagaag agggaaattc 1020
59 aatgtgacaa aagaacatag agaccaatgc acatggggat cccaaaataa gcttaccctt 1080
61 actgagggtt ctggaaaagg cacctgcata ggaaagggtc ccccatccca ccaacacctt 1140
63 tgtaaccaca ctgaagcctt taatcaaacc tctgagagtc aatatctggt acctggttat 1200
65 gacagggtgt gggcatgtaa tactggatta accccttgtg ttccactttt ggtttttaac 1260
67 caaactaaag atttttgcat tatggtccaa attgttcccc gagtgtatta ctatcccgaa 1320
69 aaagcaatcc ttgatgaata tgactacaga aatcatcgac aaaagagaga acccatatct 1380
71 ctgacacttg ctgtgatgct cggacttgga gtggcagcag gtgtaggaa aggaacagct 1440
73 gccctggtca cgggaccaca gcagctagaa acaggactta gtaacctaca tcgaattgta 1500
75 acagaagatc tccaagccct agaaaaatct gtcagtaacc tggaggaatc cctaaccctc 1560
77 ttatctgaag tagtcttaca gaatagaaga gggtttagatt tattatttct aaaagaagga 1620
79 ggattattgt tagccttgaa ggaggaatgc tgtttttatg tggatcattc aggggccatc 1680
81 agagactcca tgagcaagct tagagaaaagg ttggagaagc gtcgaaggga aaaggaaact 1740
83 actcaagggt ggtttgaggg atggttcaac aggtctcttt ggttggtctac cctactttct 1800

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,656

DATE: 01/15/2002

TIME: 07:56:34

Input Set : A:\ES.txt

Output Set: N:\CRF3\01152002\J029656.raw

```

85 gctttaacag gacccttaat agtcctcctc ctgttactca cagttgggcc atgtattatt 1860
87 aacaagttaa ttgccttcat tagagaacga ataagtgcag tccagatcat ggtacttaga 1920
89 caacagtacc aaagcccgtc tagcagggaa gctggccgc 1959
92 <210> SEQ ID NO: 2
93 <211> LENGTH: 653
94 <212> TYPE: PRT
95 <213> ORGANISM: Viral
97 <400> SEQUENCE: 2
99 Met His Pro Thr Leu Ser Arg Arg His Leu Pro Ile Arg Gly Gly Lys
100 1 5 10 15
102 Pro Lys Arg Leu Lys Ile Pro Leu Ser Phe Ala Ser Ile Ala Trp Phe
103 20 25 30
105 Leu Thr Leu Ser Ile Thr Pro Gln Val Asn Gly Lys Arg Leu Val Asn
106 35 40 45
108 Ser Pro Asn Ser His Lys Pro Leu Ser Leu Thr Trp Leu Leu Thr Asp
109 50 55 60
111 Ser Gly Thr Gly Ile Asn Ile Asn Ser Thr Gln Gly Glu Ala Pro Leu
112 65 70 75 80
114 Gly Thr Trp Trp Pro Glu Leu Tyr Val Cys Leu Arg Ser Val Ile Pro
115 85 90 95
117 Gly Leu Asn Asp Gln Ala Thr Pro Pro Asp Val Leu Arg Ala Tyr Gly
118 100 105 110
120 Phe Tyr Val Cys Pro Gly Pro Pro Asn Asn Glu Glu Tyr Cys Gly Asn
121 115 120 125
123 Pro Gln Asp Phe Phe Cys Lys Gln Trp Ser Cys Val Thr Ser Asn Asp
124 130 135 140
126 Gly Asn Trp Lys Trp Pro Val Ser Gln Gln Asp Arg Val Ser Tyr Ser
127 145 150 155 160
129 Phe Val Asn Asn Pro Thr Ser Tyr Asn Gln Phe Asn Tyr Gly His Gly
130 165 170 175
132 Arg Trp Lys Asp Trp Gln Gln Arg Val Gln Lys Asp Val Arg Asn Lys
133 180 185 190
135 Gln Ile Ser Cys His Ser Leu Asp Leu Asp Tyr Leu Lys Ile Ser Phe
136 195 200 205
138 Thr Glu Lys Gly Lys Gln Glu Asn Ile Leu Lys Trp Val His Gly Met
139 210 215 220
141 Ser Trp Gly Met Val Tyr Tyr Gly Gly Ser Gly Lys Gln Pro Gly Ser
142 225 230 235 240
144 Ile Leu Thr Ile Arg Leu Lys Ile Asn Gln Leu Glu Pro Pro Met Ala
145 245 250 255
147 Ile Gly Pro Asn Thr Val Leu Thr Gly Gln Arg Pro Pro Thr Gln Gly
148 260 265 270
150 Pro Gly Thr Ser Ser Asn Ile Thr Ser Gly Ser Asp Pro Thr Glu Ser
151 275 280 285
153 Asn Ser Thr Thr Lys Met Gly Ala Lys Leu Phe Ser Leu Ile Gln Gly
154 290 295 300
156 Ala Phe Gln Ala Leu Asn Ser Thr Thr Pro Glu Ala Thr Ser Ser Cys
157 305 310 315 320
159 Trp Leu Cys Leu Ala Leu Gly Pro Pro Tyr Tyr Glu Gly Met Ala Arg

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,656

DATE: 01/15/2002

TIME: 07:56:34

Input Set : A:\ES.txt

Output Set: N:\CRF3\01152002\J029656.raw

```

160          325          330          335
162 Arg Gly Lys Phe Asn Val Thr Lys Glu His Arg Asp Gln Cys Thr Trp
163          340          345          350
165 Gly Ser Gln Asn Lys Leu Thr Leu Thr Glu Val Ser Gly Lys Gly Thr
166          355          360          365
168 Cys Ile Gly Lys Val Pro Pro Ser His Gln His Leu Cys Asn His Thr
169          370          375          380
171 Glu Ala Phe Asn Gln Thr Ser Glu Ser Gln Tyr Leu Val Pro Gly Tyr
172 385          390          395          400
174 Asp Arg Trp Trp Ala Cys Asn Thr Gly Leu Thr Pro Cys Val Ser Thr
175          405          410          415
177 Leu Val Phe Asn Gln Thr Lys Asp Phe Cys Ile Met Val Gln Ile Val
178          420          425          430
180 Pro Arg Val Tyr Tyr Tyr Pro Glu Lys Ala Ile Leu Asp Glu Tyr Asp
181          435          440          445
183 Tyr Arg Asn His Arg Gln Lys Arg Glu Pro Ile Ser Leu Thr Leu Ala
184          450          455          460
186 Val Met Leu Gly Leu Gly Val Ala Ala Gly Val Gly Thr Gly Thr Ala
187 465          470          475          480
189 Ala Leu Val Thr Gly Pro Gln Gln Leu Glu Thr Gly Leu Ser Asn Leu
190          485          490          495
192 His Arg Ile Val Thr Glu Asp Leu Gln Ala Leu Glu Lys Ser Val Ser
193          500          505          510
195 Asn Leu Glu Glu Ser Leu Thr Ser Leu Ser Glu Val Val Leu Gln Asn
196          515          520          525
198 Arg Arg Gly Leu Asp Leu Leu Phe Leu Lys Glu Gly Gly Leu Cys Val
199          530          535          540
201 Ala Leu Lys Glu Glu Cys Cys Phe Tyr Val Asp His Ser Gly Ala Ile
202 545          550          555          560
204 Arg Asp Ser Met Ser Lys Leu Arg Glu Arg Leu Glu Lys Arg Arg Arg
205          565          570          575
207 Glu Lys Glu Thr Thr Gln Gly Trp Phe Glu Gly Trp Phe Asn Arg Ser
208          580          585          590
210 Leu Trp Leu Ala Thr Leu Leu Ser Ala Leu Thr Gly Pro Leu Ile Val
211          595          600          605
213 Leu Leu Leu Leu Leu Thr Val Gly Pro Cys Ile Ile Asn Lys Leu Ile
214          610          615          620
216 Ala Phe Ile Arg Glu Arg Ile Ser Ala Val Gln Ile Met Val Leu Arg
217 625          630          635          640
219 Gln Gln Tyr Gln Ser Pro Ser Ser Arg Glu Ala Gly Arg
220          645          650
222 <210> SEQ ID NO: 3
223 <211> LENGTH: 1959
224 <212> TYPE: DNA
225 <213> ORGANISM: Viral
227 <400> SEQUENCE: 3
228 atgcacccca cgttaagccg ggcacacctc ccgattcggg gtggaaagcc gaaaagactg      60
230 aaaatcccoct taagcttcgc ctccatcgcg tggttcctta ctctgtcaat aactcctcaa      120
232 gttaatggta aacgccttgt ggacagcccg aactcccata aacccttata tctcacctgg      180

```

RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/10/029,656

TIME: 07:56:34

Input Set : A:\ES.txt

Output Set: N:\CRF3\01152002\J029656.raw

```

234 ttacttactg actccggtac aggtattaat attaacagca ctcaagggga ggctcccttg 240
236 gggacctggt ggcctgaatt atatgtctgc cttcgatcag taatccctgg tctcaatgac 300
238 caggccacac ccccgatgt actccgtgct tacgggtttt acgtttgccc aggaccccca 360
240 aataatgaag aatattgtgg aaatcctcag gatttctttt gcaagcaatg gagctgcgta 420
242 acttctaattg atgggaattg gaaatggcca gtctctcagc aagacagagt aagttactct 480
244 tttgttaaca atcctaccag ttataatcaa tttaattatg gccatgggag atggaaagat 540
246 tggcaacagc gggtaaaaa agatgtacga aataagcaaa taagctgtca ttcgttagac 600
248 ctagattact taaaaataag tttcactgag aaaggaaaaac aagaaaatat cctaaaatgg 660
250 gtaaattgta tgtcttgggg aatggtatat tatggaggct cgggtaaaca accaggctcc 720
252 attctaacta ttcgcctcaa aataaaccag ctggagcctc caatggctat aggaccaaat 780
254 acggtcttga cgggtcaaag acccccaacc caaggaccag gaccatcctc taacataact 840
256 tctggatcag accccactga gtctaacagc acgactaaaa tgggggcaaa acttttttagc 900
258 ctcattccagg gagcttttca agctcttaac tccacgactc cagaggctac ctcttcttgt 960
260 tggctatgct tagctttggg cccaccttac tatgaaggaa tggctagaag agggaaattc 1020
262 aatgtgacaa aaggacatag agacccatgc acatggggat cccaaaataa gcttaccctt 1080
264 actgaggttt ttggaaaagg cacctgcata ggaaagggtc ccccatccca ccaacacctt 1140
266 tgtaaccaca ctgaagcctt taatcgaacc tctgagagtc agtatctggt acctggttat 1200
268 gacagggtgtt gggcatgtaa tactggatta accccttgtg tttccacttt ggtttttaac 1260
270 caaactaaag atttttgcat tatggtccaa attgttcccc gagtgtatta ctatcccga 1320
272 aaagcaatcc ttgatgaata tgactacaga aatcatcgac aaaagagaga acccatatct 1380
274 ctgacacttg ctgtgatgct cggacttgga gtggcagcag gtgtaggaac aggaacagct 1440
276 gccctgggtc cgggaccaca gcagctagaa acaggactta gtaacctaca tcgaattgta 1500
278 acagaagatc tccaagccct agaaaaatct gtcagtaacc tggaggaatc cctaacctcc 1560
280 ttatctgaag tagtcctaca gaatagaaga gggtagattt tattatttct aaaagaagga 1620
282 ggattatgtg tagccttgaa ggaggaatgc tgtttttatg tggatcattc aggggccatc 1680
284 agagactcca tgagcaagct tagagaaagg ttggagaagc gtcgaaggga aaaggaaact 1740
286 actcaagggg ggtttgagg atggttcaac aggtctcctt ggttggctac cctactttct 1800
288 gctttaacag gacccttaat agtctcctc ctgttactca cagttgggcc atgtattatt 1860
290 aacaagttaa ttgccttcat tagagaacga ataagtgcag tccagatcat ggtacttaga 1920
292 caacagtacc aaagcccgtc tagcaggga gctggccgc 1959

```

295 <210> SEQ ID NO: 4

296 <211> LENGTH: 653

297 <212> TYPE: PRT

298 <213> ORGANISM: Viral

300 <400> SEQUENCE: 4

```

302 Met His Pro Thr Leu Ser Arg Arg His Leu Pro Ile Arg Gly Gly Lys
303 1 5 10 15
305 Pro Lys Arg Leu Lys Ile Pro Leu Ser Phe Ala Ser Ile Ala Trp Phe
306 20 25 30
308 Leu Thr Leu Ser Ile Thr Pro Gln Val Asn Gly Lys Arg Leu Val Asp
309 35 40 45
311 Ser Pro Asn Ser His Lys Pro Leu Ser Leu Thr Trp Leu Leu Thr Asp
312 50 55 60
314 Ser Gly Thr Gly Ile Asn Ile Asn Ser Thr Gln Gly Glu Ala Pro Leu
315 65 70 75 80
317 Gly Thr Trp Trp Pro Glu Leu Tyr Val Cys Leu Arg Ser Val Ile Pro
318 85 90 95
320 Gly Leu Asn Asp Gln Ala Thr Pro Pro Asp Val Leu Arg Ala Tyr Gly
321 100 105 110

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,656

DATE: 01/15/2002

TIME: 07:56:34

Input Set : A:\ES.txt

Output Set: N:\CRF3\01152002\J029656.raw

```

323 Phe Tyr Val Cys Pro Gly Pro Pro Asn Asn Glu Glu Tyr Cys Gly Asn
324      115      120      125
326 Pro Gln Asp Phe Phe Cys Lys Gln Trp Ser Cys Val Thr Ser Asn Asp
327      130      135      140
329 Gly Asn Trp Lys Trp Pro Val Ser Gln Gln Asp Arg Val Ser Tyr Ser
330 145      150      155      160
332 Phe Val Asn Asn Pro Thr Ser Tyr Asn Gln Phe Asn Tyr Gly His Gly
333      165      170      175
335 Arg Trp Lys Asp Trp Gln Gln Arg Val Gln Lys Asp Val Arg Asn Lys
336      180      185      190
338 Gln Ile Ser Cys His Ser Leu Asp Leu Asp Tyr Leu Lys Ile Ser Phe
339      195      200      205
341 Thr Glu Lys Gly Lys Gln Glu Asn Ile Leu Lys Trp Val Asn Gly Met
342      210      215      220
344 Ser Trp Gly Met Val Tyr Tyr Gly Gly Ser Gly Lys Gln Pro Gly Ser
345 225      230      235      240
347 Ile Leu Thr Ile Arg Leu Lys Ile Asn Gln Leu Glu Pro Pro Met Ala
348      245      250      255
350 Ile Gly Pro Asn Thr Val Leu Thr Gly Gln Arg Pro Pro Thr Gln Gly
351      260      265      270
353 Pro Gly Pro Ser Ser Asn Ile Thr Ser Gly Ser Asp Pro Thr Glu Ser
354      275      280      285
356 Asn Ser Thr Thr Lys Met Gly Ala Lys Leu Phe Ser Leu Ile Gln Gly
357      290      295      300
359 Ala Phe Gln Ala Leu Asn Ser Thr Thr Pro Glu Ala Thr Ser Ser Cys
360 305      310      315      320
362 Trp Leu Cys Leu Ala Leu Gly Pro Pro Tyr Tyr Glu Gly Met Ala Arg
363      325      330      335
365 Arg Gly Lys Phe Asn Val Thr Lys Gly His Arg Asp Pro Cys Thr Trp
366      340      345      350
368 Gly Ser Gln Asn Lys Leu Thr Leu Thr Glu Val Phe Gly Lys Gly Thr
369      355      360      365
371 Cys Ile Gly Lys Val Pro Pro Ser His Gln His Leu Cys Asn His Thr
372      370      375      380
374 Glu Ala Phe Asn Arg Thr Ser Glu Ser Gln Tyr Leu Val Pro Gly Tyr
375 385      390      395      400
377 Asp Arg Trp Trp Ala Cys Asn Thr Gly Leu Thr Pro Cys Val Ser Thr
378      405      410      415
380 Leu Val Phe Asn Gln Thr Lys Asp Phe Cys Ile Met Val Gln Ile Val
381      420      425      430
383 Pro Arg Val Tyr Tyr Tyr Pro Glu Lys Ala Ile Leu Asp Glu Tyr Asp
384      435      440      445
386 Tyr Arg Asn His Arg Gln Lys Arg Glu Pro Ile Ser Leu Thr Leu Ala
387      450      455      460
389 Val Met Leu Gly Leu Gly Val Ala Ala Gly Val Gly Thr Gly Thr Ala
390 465      470      475      480
392 Ala Leu Val Thr Gly Pro Gln Gln Leu Glu Thr Gly Leu Ser Asn Leu
393      485      490      495
395 His Arg Ile Val Thr Glu Asp Leu Gln Ala Leu Glu Lys Ser Val Ser

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/029,656

DATE: 01/15/2002

TIME: 07:56:35

Input Set : A:\ES.txt

Output Set: N:\CRF3\01152002\J029656.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:428 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:440 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:452 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:464 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:1084 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17
L:1178 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:18
L:1190 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:19
L:1202 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:20
L:1214 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:21
L:1226 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22
L:1238 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23